Problem Statement 1

The college bookstore tells prospective students that the average cost of its textbooks is Rs.52 with a standard deviation of Rs. 4.50. A group of smart statistics students thinks that the average cost is higher. To test the bookstore's claim against their alternative, the students will select a random sample of size 100. Assume that the mean from their random sample is Rs. 52.80. Perform a hypothesis test at the 5% level of significance and state your decision.

Solution 1:

$$H_0: \mu = Rs.52 \quad vs \quad H_1: \mu > Rs.52$$

$$n = 100, \quad \bar{x} = Rs.52.80, \quad \sigma = Rs.4.50, \quad \alpha = 0.05$$

calculate the Z score as follows

$$Z^* = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{52.8 - 52}{\frac{4.50}{\sqrt{100}}} = \frac{0.8}{0.45} = \mathbf{1.778}$$

Now we need to compute the appropriate ***p-value*** based on our alternative hypothesis.

$$P(Z > Z^*) = P(Z > 1.778)$$

$$= 1 - P(Z < 1.778)$$

$$= 1 - 0.9623 = \mathbf{0.0377} < 0.05$$

Since the calculated *p-value* (**0.0377**) is lower than the level of significance (**0.05**), we reject the null hypothesis in favour of the alternative hypothesis and draw the same conclusion as above that the average cost of the bookstore textbook is higher than Rs. 52

Problem Statement 2

A certain chemical pollutant in the Genesee River has been constant for several years with mean μ = 34 ppm (parts per million) and standard deviation σ = 8 ppm. A group of factory representatives whose companies discharge liquids into the river is now claiming that they have lowered the average with improved filtration devices. A group of environmentalists will test to see if this is true at the 1% level of significance. Assume that their sample of size 50 gives a mean of 32.5 ppm. Perform a hypothesis test at the 1% level of significance and state your decision.

Solution 2:

$$H_0: \mu = 34\ ppm \quad vs \quad H_1: \mu < 34\ ppm$$

$$n = 50, \quad \bar{x} = 32.5\ ppm, \quad \sigma = 8\ ppm, \quad \alpha = 0.01$$

Using the Z score method we have:

$$Z^* = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{32.5 - 34}{\frac{8}{\sqrt{50}}} = \frac{-1.5}{1.1314} = -1.3258$$

Now we need to compute the appropriate ***p-value*** based on our alternative hypothesis.

$$P(Z < Z^*) = P(Z < -1.3258)$$

$$= P(Z > 1.3258)$$

$$= 0.0925 > 0.01$$

Since the calculated *p-value* (**0.0925**) is higher than the level of significance (**0.01**), we fail to reject the null hypothesis in favour of the alternative hypothesis and conclude that the mean is unchanged even after use of improved filtration devices.

Problem Statement 3

Based on population figures and other general information on the U.S. population, suppose it has been estimated that, on average, a family of four in the U.S. spends about $1135 annually on dental expenditures. Suppose further that a regional dental association wants to test to determine if this figure is accurate for their area of country.

To test this, 22 families of 4 are randomly selected from the population in that area of the country and a log is kept of the family's dental expenditure for one year. The resulting data are given below. Assuming, that dental expenditure is normally distributed in the population, use the data and an alpha of **0.05** to test the dental association's hypothesis.
1008, 812, 1117, 1323, 1308, 1415, 831, 1021, 1287, 851, 930, 730, 699, 872, 913, 944, 954, 987, 1695, 995, 1003, 994

Solution 3:

$$\bar{x} = \frac{1}{22}\sum X_i = \frac{1}{22}(1008 + 812 \dots + 994) = \mathbf{1031.32}$$

$$\sigma^2 = \frac{1}{22-1}\sum (X_i - \bar{X})^2 = \frac{1}{22-1}\left[(1008 - 1031.32)^2 + \cdots + (994 - 1031.32)^2\right] = \mathbf{57\,779.54}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{57\,779.54} = \mathbf{240.37}$$

Thus

$$n = 22, \qquad \bar{x} = 1031.32, \qquad \sigma = 240.37\,, \qquad \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$$

Therefore to determine if the mean figure is accurate for the area of the country, we test the hypothesis and this is a two tailed test

$$H_0: \mu = \$\,1135 \quad vs \quad H_1: \mu \neq \$1135$$

The critical value of Z at 5% level of significance if $Z = \pm 1.96$

$$Z^* = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{1031.32 - 1135}{\frac{240.37}{\sqrt{22}}} = \frac{-103.68}{51.25} = -2.0231$$

This value lies in the rejection region

Now we compute the appropriate ***p-value*** based on our alternative hypothesis.

$$P(-2.0231 < Z < 2.0231) = 2 * P(Z < 2.0231) - 1$$

$$= 2 * 0.9785 - 1 = \mathbf{0.9569} > 0.05$$

We reject the null hypothesis and conclude that for their area of the country, the mean dental expenditure is different from the country's population mean, hence the general U.S. population figure is not accurate for their region.

Problem Statement 4:

In a report prepared by the Economic Research Department of a major bank the Department manager maintains that the average annual family income on Metropolis is $48,432. What do you conclude about the validity of the report if a random sample of 400 families shows and average income of $48,574 with a standard deviation of 2000?

Solution 4:

$$n = 400, \qquad \bar{x} = 48574, \qquad \sigma = 2000, \qquad \alpha = 0.05$$

This is large sample therefore the central limit theorem holds

We reject H_0 at 5% level of significance if $Z < -1.96$ or $Z > +1.96$

$$H_0: \mu = \$\,48{,}432 \quad vs \quad H_1: \mu > \$48{,}432$$

$$Z^* = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{48574 - 48432}{\frac{2000}{\sqrt{400}}} = \frac{142}{100} = 1.42$$

Now we compute the appropriate ***p-value*** based on our alternative hypothesis.

$$P(Z < -1.42) + P(Z > 1.42) = P(Z < -1.42)1 + P(Z < 1.42)$$

$$= \mathbf{0.1556} > 0.05$$

We cannot reject the null hypothesis at 5% level of significance, therefore we conclude that the average family income in Metropolis is 48,432.

Problem Statement 5

Suppose that in past years the average price per square foot for warehouses in the United States has been $32.28. A national real estate investor wants to determine whether that figure has changed now. The investor hires a researcher who randomly samples 19 warehouses that are for sale across the United States and finds that the mean price per square foot is $31.67, with a standard deviation of $1.29. Assume that the prices of warehouse footage are normally distributed in population. If the researcher uses a 5% level of significance, what statistical conclusion can be reached? What are the hypotheses?

Solution 5:

$$n = 19, \quad \bar{x} = 31.67, \quad \sigma = 1.29, \quad \alpha = 0.05$$

We reject H_0 at 5% level of significance if $Z < -1.96$ or $Z > +1.96$

$$H_0: \mu = 32.28 \quad vs \quad H_1: \mu \neq 32.28$$

This is a two tailed test and the critical values of Z are $Z = -1.96$ and $Z = +1.96$

$$Z^* = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{31.67 - 32.28}{\frac{1.29}{\sqrt{19}}} = \frac{-0.61}{0.29}$$

$$= \mathbf{-2.0612}$$

This figure falls in the rejection region, we therefore reject the null hypothesis and conclude that the average price per square foot for the warehouses has changed

Problem Statement 6:

Fill in the blank spaces in the table and draw your conclusions from it. Take σ = 2.5, assume all other necessary parameters.

Acceptance Region	Sample Size	α	Β at μ = 52	Β at μ = 50.5
48.5 < x < 51.5	10	0.0576	0.2643	0.8923
48 < x < 52	10	0.0114		
48.81 < x < 51.19	16			
48.42 < x < 51.58	16			

Solution 6:

Solution 6

48 < x < 52 n=10

Β at μ = 52

$$= \frac{48-52}{\frac{2.5}{\sqrt{10}}} = \frac{-4}{0.791} = -5.06$$

$$= \frac{52-52}{\frac{2.5}{\sqrt{10}}} = \frac{0}{0.791} = 0$$

$$= P(-5.06 \leq Z \leq 0)$$

$$= P(Z \leq 0) - P(Z \leq -5.06)$$

= 0.5-0

=0.5

$48 < x < 52$ n=10

B at μ = 50.5

$$=\frac{48-50.5}{\frac{2.5}{\sqrt{10}}}=\frac{-2.5}{0.791}=-3.16$$

$$=\frac{52-50.5}{\frac{2.5}{\sqrt{10}}}=\frac{1.5}{0.791}=1.90$$

$$=P(-3.16 \le Z \le 1.9)$$

$$=P(Z \le 1.9)-P(Z \le -3.16)$$

= 0.9713–(1-0.9992)

=0.9705

$48.81 < x < 51.19$

n= 16

B at μ = 52

$$=\frac{48.81-52}{\frac{2.5}{\sqrt{16}}}=\frac{-3.19}{0.625}=-\mathbf{5.104}$$

$$=\frac{51.19-52}{\frac{2.5}{\sqrt{16}}}=\frac{-0.81}{0.625}=-\mathbf{1.30}$$

$$=P(-5.10 \le Z \le -1.30)$$

$$=P(Z \le -1.30)-P(Z \le -5.1)$$

= (1-0.9032)-0

=0.0968

B at μ = 50.5

$$=\frac{48.81-50.5}{\frac{2.5}{\sqrt{16}}}=\frac{-1.69}{0.625}=-\mathbf{2.7}$$

$$=\frac{51.19-50.5}{\frac{2.5}{\sqrt{16}}}=\frac{-0.69}{0.625}=-\mathbf{1.1}$$

$$=P(-2.7 \le Z \le -1.1)$$

$$=P(Z \le -1.1)-P(Z \le -2.7)$$

= (1-0.8643)- (1-0.9965)

=0.1322

$48.42 < x < 51.58$

n= 16

B at μ = 52

$$=\frac{48.42-52}{\frac{2.5}{\sqrt{16}}}=\frac{-3.58}{0.625}=-\mathbf{5.73}$$

$$=\frac{51.58-52}{\frac{2.5}{\sqrt{16}}}=\frac{-0.42}{0.625}=-0.67$$

$$=P(-5.73\leq Z\leq -0.67)$$

$$=P(Z\leq -0.67)-P(Z\leq -5.73)$$

= (1-0.7486)-0

=0.2514

B at μ = 50.5

$$=\frac{48.42-50.5}{\frac{2.5}{\sqrt{16}}}=\frac{-2.08}{0.625}=-3.33$$

$$=\frac{51.58-50.5}{\frac{2.5}{\sqrt{16}}}=\frac{1.08}{0.625}=\mathbf{1.73}$$

$$=P(-3.33\leq Z\leq 1.73)$$

$$=P(Z\leq 1.73)-P(Z\leq -3.33)$$

= (1-0.9582)- (1-0.9996)

=0.0414

α

$48.81 < x < 51.19$

at μ = 50

n=16

$$=\frac{48.81-50}{\frac{2.5}{\sqrt{16}}}=\frac{-1.19}{0.625}=-1.904$$

$$=\frac{51.19-50}{\frac{2.5}{\sqrt{16}}}=\frac{1.19}{0.625}=1.904$$

$$=P(Z<-1.904)+P(Z>1.904)$$

=(1-0.9713)+ (1-0.9713)

=0.0574

α

$48.42 < x < 51.58$

at μ = 50

n=16

$$=\frac{48.42-50}{\frac{2.5}{\sqrt{16}}}=\frac{-1.58}{0.625}=-2.528$$

$$=\frac{51.58-50}{\frac{2.5}{\sqrt{16}}}=\frac{1.58}{0.625}=2.528$$

$$=P(Z < -2.528) + P(Z > 2.528)$$

$$=(1-0.9943)+ (1-.9943)$$

$$=0.0114$$

Acceptance Region	Sample Size	α	Β at μ = 52	Β at μ = 50.5
48.5 < x < 51.5	10	0.0576	0.2643	0.8923
48 < x < 52	10	0.0114	0.5	0.9705
48.81 < x < 51.19	16	0.0574	0.0968	0.1322
48.42 < x < 51.58	16	0.0114	0.2514	0.0414

Problem Statement 7

Find the t-score for a sample size of 16 taken from a population with mean 10 when the sample mean is 12 and the sample standard deviation is 1.5.

Solution 7:

$$n = 16, \quad \bar{x} = 12, \quad s = 1.5, \quad \mu = 10$$

$$t^* = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{12 - 10}{\frac{1.5}{\sqrt{16}}} = \frac{16}{3} = \mathbf{5.333}$$

Problem Statement 8

Find the t-score below which we can expect 99% of sample means will fall if samples of size 16 are taken from a normally distributed population.

Solution 8:

$$1 - \alpha = 0.99 \ \rightarrow \alpha = 1 - 0.99 = 0.01$$

$$df = n - 1 = 16 - 1 = 15$$

$$t^*_{0.99}(15) = \ -t^*_{0.01}(15) = -2.602$$

Problem Statement 9

A restaurant has been arranging sales of 300 lunch packets per day at Brigade road. Because of the construction of the new building and other complexes, it expects to increase the sale. During the first 16 days after the occupation of these building, the daily sales were 304, 367, 385, 386, 262, 329, 302, 292, 350, 320, 298, 258, 364, 294, 276, 333. Based on this information will you conclude that the sales have increased?

Solution 8:

$$\bar{x} = \frac{1}{16}\sum X_i = \frac{1}{16}(304 + 367 + \cdots + 333) = \mathbf{320}$$

$$s^2 = \frac{1}{16-1}\sum (X_i - \bar{X})^2 = \frac{1}{15}\ [(304 - 320)^2 + \cdots + (333 - 320)^2] = \mathbf{1702.93}$$

$$s = \ \sqrt{s^2} = \sqrt{1702.93} = \mathbf{41.27}$$

Thus $\quad n = 16, \quad \bar{x} = 320, \quad s = 41.27, \quad \mu = 300, \quad \propto = 0.05$

$$H_0: \mu = 300 \quad vs \quad H_1: \mu > 300$$

We reject H_0 at 5% level of significance if $Z < \ -1.96$ or $Z > +1.96$

$$Z^* = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$= \frac{320 - 300}{\frac{41.27}{\sqrt{16}}} = \frac{20}{10.3175} = \mathbf{1.9384}$$

The appropriate ***p-value*** is given by

$$P(Z > Z^*) = P(Z > 1.9384)$$

$$= 1 - P(Z < 1.9384)$$

$$= 1 - 0.9737 = \mathbf{0.0263} < 0.05$$

The ***z-score*** falls within the acceptance region and the ***p-value*** is less than the level of significance which makes us ***fail to reject*** the NULL hypothesis and conclude that the sales have not increased.